

doi: 10.7690/bgzdh.2013.05.025

一种面向物化视图的远程增量维护系统

赵建平¹, 周锦标¹, 何剑伟¹, 丁冉¹, 赵建辉²

(1. 中国卫星海上测控部技术部, 江苏 江阴 214431; 2. 空间信息中继传输技术研究中心, 北京 100094)

摘要: 为有效保证客户端物化视图与数据源的一致性并降低网络开销, 提出一种物化视图增量维护方法。采用 XML 物化视图远程增量维护方法, 通过系统原型设计、系统部署、同步机制, 实现试验任务数据接口同步。结果表明: 在不影响试验任务数据处理与传输的实时性及可靠性的条件下, XML 物化视图远程增量维护系统可有效实现指挥显示系统数据接口的一致性、可靠性和安全性。

关键词: 物化视图; 远程增量维护; 数据接口; 同步

中图分类号: TJ86 **文献标志码:** A

A Web Incremental Maintenance System Oriented Materialized View

Zhao Jianping¹, Zhou Jinbiao¹, He Jianwei¹, Ding Ran¹, Zhao Jianhui²

(1. *Technology Department, Satellite Maritime Tracking & Controlling Department of China, Jiangyin 214431, China;*
2. *Spatial Information Relay Transmission Technology Research Center, Beijing 100094, China*)

Abstract: To ensure the materialized views on the client sides consistent with the data source and reduce the network cost, a kind of incremental materialized view maintenance method is put forwards. Using the XML materialized view web incremental maintenance method, through the system prototype design, system deployment and synchronization mechanism, the experiment task data interface synchronization can be realized. The results show that, in the condition of not affecting the real-time performance and the reliability of the experiment data processing and transmitting, XML materialized view web incremental maintenance system can effectively achieve the command and display system data interface's consistency, reliability and security.

Key words: materialized view; web incremental maintenance; data interface; synchronization

0 引言

物化视图既与普通视图一样存储底层查询定义, 又存储由底层查询产生的结果集并占用数据库的存储空间。物化视图存储基于远程表的数据, 称为快照。对于实际存储数据, 物化视图的处理速度比普通视图快。XML 表示的查询视图, 因其可扩展性和灵活性, 被广泛应用于 WEB 应用系统, 是网络上通用的数据传输格式。通过定义 XML 架构和数据库中的表之间的映射, 可以创建持久性数据的“XML 视图”^[1]。

在试验任务数据库中, 将其数据视图进行 XML 格式的存储, 形成基于 XPath 或 XQuery 语言的物化视图, 可提高试验任务数据库的查询效率和灵活性^[1]; 但是 XML 物化视图也为试验任务 IP 网络高实时性、大数据量、高可靠性的数据传输带来了一定影响: 1) 由于物化视图是实实在在的数据区域, 要占据存储空间; 2) 当物化视图底层的数据源变化时, 为了保持数据的一致性, 系统需要刷新(更新)物化视图, 也会耗费一定的资源。基于上述原因, 对试验任务数据库进行远程操作时, 其高信息量、复杂的数据处理, 会给试验 IP 网及服务器和客户端

造成资源负担; 因此, 笔者提出一种物化视图增量维护方法, 以实现系统查询过程的优化和数据视图的快速更新与访问。

1 增量维护原理

1.1 物化视图增量维护

当客户端物化视图所映射的数据库服务器的数据源变化时, 尽管微量的数据源变化只影响了部分视图数据, 但是对快照的数据区域全集的更新维护, 会因重新计算视图而导致较大的网络资源开销。物化视图增量维护, 根据视图变化的部分来更新视图, 只重新计算底层数据库中的变化部分, 可减少重复查询的开销。当底层数据库的内容发生变化时, 通知数据库系统, 数据库系统根据底层数据库的变化情况和物化视图的定义, 求出物化视图的增量, 然后相应地修改物化视图。

1.2 原理

假设数据源为 x , 客户端请求的查询计算为 f , 其查询计算结果为 r , 则

$$r=f(x)^{[2]} \quad (1)$$

在树型结构的 XML 数据源中, 其变更是以节

收稿日期: 2012-11-04; 修回日期: 2012-12-07

作者简介: 赵建平(1974—), 女, 福建人, 硕士, 从事软件工程及计算机应用研究。

点为最小单位的，任何 XML 数据源的变更都可以归纳为对某个指定节点的修改、插入、删除操作。

数据源所发生的变更为 Δx ，为了计算 $f(x+\Delta x)$ 的计算结果 r' ，应找到对应于 f 的增量计算 f_{inc} ，使 $r'=f_{inc}(x, \Delta x, r)^{[2]}$ (2)

在增量计算 f_{inc} 的设计中，计算的过程中可以尽可能利用前次的计算结果 r ，并尽可能少地访问原有的 XML 数据源 x ，那么 f_{inc} 的计算量就有可能远远小于 f 的计算量，进而实现 XML 物化视图的增量更新。笔者已经实现了根据 XPath 或 XQuery 查询 f 来组织增量计算 f_{inc} 的几种代码生成方法，可以根据 XML 视图定义语言 (XPath 或 XQuery) 的不同来设计不同的增量计算算法。

在 Web 环境中，XML 查询请求 f 从客户端发出，XML 数据源 x 存储在服务器端， $f(x)$ 的前次计算结果 r 保存在客户端。

客户服务器软件来支持增量维护计算；服务器配置于数据源一侧，管理数据源以及针对 XML 查询 f 所生成的增量维护程序 f_{inc} ；当数据源 x 发生数据变化 Δx (或者当客户查询请求的参数数值变化 Δx) 时，生成 $f_{inc}(x, \Delta x, r)$ ，并将 f_{inc} 计算分解为 2 个部分，在服务器端完成与 x 和 Δx 相关的计算，生成增量更新 f_{res} 。这个动态生成的 f_{res} 程序代码被发送到客户端专门处理前次的计算结果 r 并得到新的计算结果 r' ，即 $r'=f_{res}(r)$ 。

1.3 XQuery 技术

可扩展标记语言的查询语言 (an XML query language, XQuery) 是一种可以查询结构化或半结构化 XML 数据的语言。XQuery 基于现有的 XPath 查询语言，并支持更好的迭代、更好的排序结果以及构造必需的 XML 的功能。XQuery 在 XQuery 数据模型上运行。此模型是 XML 文档以及可能为类型化或非类型化的 XQuery 结果的抽象概念。类型信息基于 W3C XML 架构语言所提供的类型。如果没有可用的类型化信息，XQuery 将按照非类型化处理数据。这与 XPath 1.0 版处理 XML 的方式相似^[3]。

Xquery 1.0 的路径表示法：其语法是基植于 XPath 1.0 的语法，以路径方式浏览 XML 文件。路径表示法的开头可指定文件中的特定节点或一包含其他子节点的母节点，再依文件的结构配合 XPath 的语法，找到符合路径的数据^[4]。例如，要按书的编号定位书的信息节点，可以用路径表示法表示为：

```
document("file://D:\xquark\MyLibrary.xml")/
books/book[bookID="T10283"]
```

2 系统架构

远程增量维护系统采用 C/S (Client/Server, 客户/服务器) 架构，服务器提供“增量维护代理”，生成“增量更新程序”；客户端执行“增量更新程序”，对物化视图进行更新，如图 1。

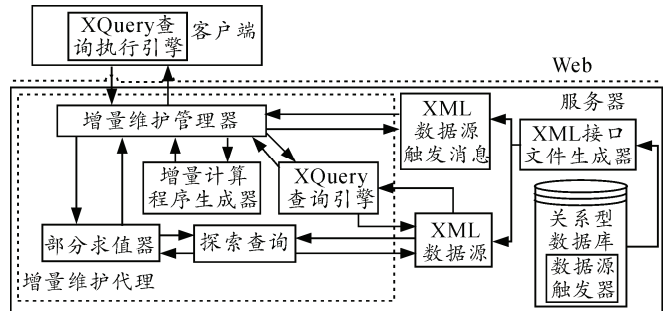


图 1 系统架构

各关键模块功能^[4]如下：

1) 客户端：“XQuery 查询引擎”+“增量更新程序执行引擎”，提供访问物化视图的应用接口，并执行从服务器返回的“增量更新程序”，对物化视图进行更新；物化视图以 XML 的格式存储于客户端。

2) 服务器端：“增量维护代理”服务，负责根据数据源的变化生成“增量更新程序”。由“增量维护管理器”、“增量计算程序生成器”、“部分求值器”、“探索查询”和“XQuery 查询引擎”等模块组成。

① “增量维护管理器”，负责调度增量维护服务的工作流程，为客户端提供增量维护服务的接口；协调客户端请求和数据源变更响应工作流程的并行处理，为客户端建立 XQuery 查询注册信息，管理“增量更新程序”的存储，维护“增量更新程序”与对应客户端的版本一致。

② XQuery 查询引擎，为客户端的首次查询获取 XML 视图提供查询接口。

③ “增量计算程序生成器”，负责“增量计算程序”的代码生成。此模块被“增量维护管理器”调用，通过输入数据源的变更信息和客户端注册的 XQuery 表达式，生成对应的“增量计算程序”代码。

④ 部分求值+探索查询，“部分求值”调用探索查询来完成涉及 XML 数据源文档的计算，并将计算结果代入“增量计算程序”进行代换，得到增量更新程序。

⑤ “数据源触发器”负责捕获发生于数据库数据源的所有变化，向“XML 接口文件生成器”提交数据源变更通知。

⑥ “接口文件生成器”负责根据数据源变更通知生成“XML 数据源触发消息”，向“增量维护管理器”提交 XML 数据源变更消息；消息包括删除某指定节点、插入节点、修改节点，“接口文件生

成器”将关系数据库表生成 XML 数据源。

3 应用实例

3.1 任务数据 XML 接口文件

试验信息监视与显示软件系统 (experiment information monitoring and display software system, EIMDSS) 的主要功能是对航天试验任务的信息进行实时监视和显示, 为任务总体人员及指挥者提供有效的数据支持和辅助决策。在 EIMDSS 中, 数据处理软件接收来自信息中心的各类信息源码, 依据数据库中信息类别表、各类信息格式表的配置进行处理解算后, 按系统内部的接口约定, 网络发送至数据显示软件, 为显示软件提供各种任务测控数据。

面对不同任务中不同数据内容及不同数据显示要求, 系统将易变化的显示页面要素封装在数据显示软件的显示控件中^[5]。数据处理软件和数据显示软件之间需要通过 XML 格式的接口文件明确数据交换的内容和格式信息, 保证数据的一致性、完整性和安全性。在试验任务数据库中, 存有各类信息帧处理方法的配置, 这些配置信息包括数据帧内容、具体处理方法等。通过这些配置信息可生成数据处理软件信息帧处理结果的内容和格式约束信息, 这就是数据处理软件与数据显示软件的接口; 将这些接口信息写入 XML 文件便是系统的内部数据接口。因此, 接口文件生成器连接任务数据库, 根据信息类别表、各类信息格式表的配置, 生成接口文件 DisplayInitXML.xml。数据显示软件被设计为编辑和运行状态。数据显示软件通过加载此接口文件, 获得数据内容和数据格式接口信息, 在编辑状态, 依据 XML 接口文件进行显示控件的数据绑定, 完成显示页面编辑功能; 在运行状态, XML 接口文件作为数据解析的依据, 完成试验信息数据的实时监视显示功能。试验信息监视与显示软件接口如图 2。

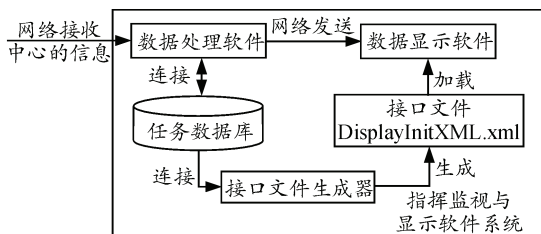


图 2 试验信息监视与显示软件接口

3.2 物化视图增量维护的部署

某次型号任务所依据的任务测控信息系统信息交换规定发生变更时, 那么任务数据库中本次任务某些表的配置就会变化, 部署了数据显示软件的 20 多台数据显示客户端所有的接口文件必须得到及时

的更新, 以保证接口的一致性。系统部署如图 3。

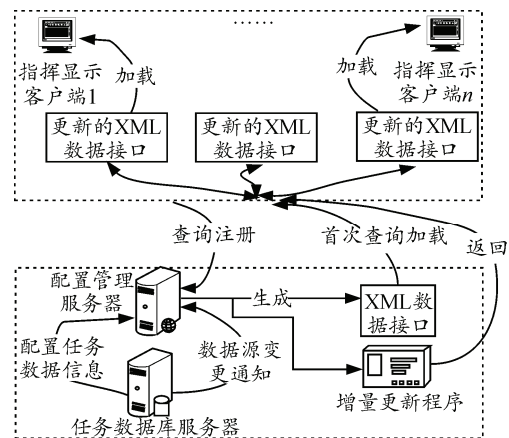


图 3 系统部署

1) 指挥显示客户端部署数据显示软件, 该软件内置“XQuery 查询求值引擎”模块, 具有访问物化视图的应用接口, 执行从服务器返回的“增量更新程序”, 更新物化视图, 将物化视图以 XML 的格式存储于客户端, 数据显示软件加载该 XML 格式的接口文件, 对显示控件的进行数据绑定, 完成接收帧统计、星下点轨迹、速高曲线、轨道分参、外测数据曲线、箭遥参数等显示页面的定制。

2) 任务数据库服务器部署 SQL Server 2005 (SQL Server 2005 包含支持 XPath 和 XQuery 的内置查询求值引擎, 支持 XML 格式的数据集处理功能) 和任务数据库 EIMDSS_DB, 历次任务中所需用到信息的类型、各种信息格式以数据表的存储于该关系型数据库中。

3) 配置管理服务器端部署“增量维护代理”模块, 同时部署“任务数据管理系统”和“接口文件生成器”。“任务数据管理系统”连接任务数据库, 对任务数据信息进行配置管理, 主要负责数据库配置人员对数据表的更新操作(因为任务数据信息的配置)。“接口文件生成器”负责根据数据源触发器的触发将关系型数据表生成 XML 数据源并生成 XML 数据源触发消息。“增量维护代理”则完成第 2 章所述功能。

3.3 物化视图增量维护的过程

通过使用物化视图增量维护的方法, 数据显示客户端加载更新数据接口文件的过程可分 3 部分:

1) 首次查询。

数据显示客户端的数据显示软件一旦(首次)启动, 便依据输入任务代号 XX-XX, 向服务器发出本次任务数据接口的查询请求:

```
let Y b:=doc("XX-XXDisplayInitXML.xml")//
return Y b
```


增量维护管理器,监听客户端发出的查询请求,首先判断是否首次查询,如果是,然后对此 XQuery 表达式进行查询工作,得到数据接口的查询结果,查询完成后,进行注册信息,最后将结果视图返回客户端,以物化视图形式存储于数据显示客户端(默认路径下),数据显示软件便加载该接口文件。

2) 数据源变化的影响。

① 一旦任务数据库数据源发生变化时,“数据源触发器”触发“接口文件生成器”通过对已有 XML 文件的访问,构造生成“XML 数据源触发消息”(触发消息包含更新节点的 XPath 表达式和更新节点内容),并生成更新后的 XML 数据源;

② “XML 数据源触发消息”向“增量维护管理器”发出变化通知,“增量维护管理器”调用“增量计算程序生成器”,并传入存储在查询注册表中的变更信息(XPath 表达式),经“增量计算程序生成器”计算,生成相应的“增量计算程序”;

③ “增量维护管理器”调用“部分求值器”,并将生成的“增量计算程序”传入,进行部分求值,得到“增量更新程序”,以队列的形式存储于配置管理服务器端。

如将 GPS 弹道数据外部信类编码 BID 改为 0x002A0106,更新语句如下:

```
REPLACE      NODE<OuterCode>002A0106</OuterCode>
```

```
WITHdoc(“d:\XX-XXDisplayInitXML.xml”)//
types [InfoName=” GPS 弹道”]/OuterCode
```

得到的增量更新程序为:

```
fn:subsequence( $ rs’1’5) ‘<OuterCode>
002A0106</OuterCode>’ fn:subsequence($ rs’7)
```

3) 非首次查询。

数据显示客户端的数据显示软件启动,再次通过输入同样的任务代号 XX-XX 向服务器发出任务数据接口的查询请求:

```
let $ b:=doc(“XX-XXDisplayInitXML.xml”)//
return $ b(数据接口刷新)
```

查询将返回相应的增量更新程序:

```
fn:subsequence( $ rs’1’5) ‘<OuterCode>
002A0106</OuterCode>’ fn:subsequence($ rs’7)
```

数据显示客户端在首次查询结果的基础上执行返回的增量更新程序,便得到最新的数据接口文件。

某一任务所有数据显示客户端的查询请求是相同的,可以进行统一维护。

3.4 物化视图增量维护同步机制

当客户端的增量维护请求与 XML 数据源触发消息同时发生时,增量维护代理需能对此进行并行

处理;当多客户端同时向增量维护代理发送请求时,增量维护代理还需并行处理多 XQuery 查询请求。同步机制如图 4。

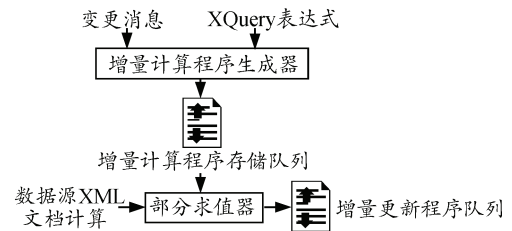


图 4 同步机制

1) 客户端请求查询的同步过程。

因可能随时接收客户端的查询请求,增量维护代理接收请求和执行查询的过程往往是同步进行的,所以在增量维护代理中设置数据的共同访问区(注册及更新程序列表)。

增量维护管理器,监听客户端发出的查询请求,首先判断是否首次查询,如果是首次查询,然后对此 XQuery 表达式进行查询工作,查询完成后,进行注册,最后将结果视图返回客户端。

2) 数据源更新通知的同步过程。

当 XML 数据源触发消息发送给增量维护管理器 XML 数据源更新消息,增量维护管理器将更新通知加入变更消息队列中。从变更消息队列中获取变更信息,同时访问注册及更新列表,依次生成更新程序,并将更新程序加入注册及更新程序列表。

4 结束语

基于 XQuery 语言的物化视图提高了试验任务数据库的查询效率和灵活性。物化视图增量维护实现了查询过程的优化和数据视图的快速更新,既降低了网络开销和服务器的负担,又保证了试验任务数据处理与传输的实时性与可靠性,保证了数据显示客户端物化视图与任务数据库数据源的一致性,确保了指挥显示监视软件系统内部数据接口的一致性、可靠性和安全性。

参考文献:

[1] 万常选. XML 数据库技术[M]. 北京:清华大学出版社, 2005: 20.
 [2] 孟彦, 廖湖声, 金雪云, 等. 面向 XML 物化视图远程增量维护的版本管理技术研究[J]. 计算机工程与设计, 2008, 29(19): 5054-5058.
 [3] 陈荣鑫. 基于 XQuery 的 XML 应用系统集成[J]. 山东理工大学学报: 自然科学版, 2011(25): 11-15.
 [4] 彭蕾, 廖湖声, 金雪云. XQuery 物化视图增量更新系统框架的研究[J]. 计算机应用与软件, 2011, 28(6): 138-142.
 [5] 段慧芬, 孙丰, 薛倡新. 测站试验任务软件接口设计与实现[J]. 飞行器测控学报, 2009(28): 55-59.